HESSD

Interactive
comment

Interactive comment on "A comprehensive sensitivity and uncertainty analysis for discharge and nitrate-nitrogen loads involving multiple discrete model inputs under future changing conditions" *by* Christoph Schürz et al.

Anonymous Referee #3

Received and published: 3 October 2018

I provide my comments below according to the HESS review criteria. Given some of my major comments below, it does not seem necessary to provide a more detailed line by line annotation at this point.

1. Does the paper address relevant scientific questions within the scope of HESS? Yes. Trying to quantify and attribute uncertainty from various sources in "eco-hydrological" modelling in the context of climate and environmental change.

2. Does the paper present novel concepts, ideas, tools, or data? I found the way

figures 2 to 7 very informative. I particularly found figure 2 very appealing in presenting SA results.

3. Are substantial conclusions reached? Given some of the discussions provided on the methodology below in NO.4, I am not sure if we can say conclusions are substantial.

4. Are the scientific methods and assumptions valid and clearly outlined? I very much liked how the manuscript tries to do a systematic and comprehensive approach, step-by-step, to set up the models, define scenarios, conduct SA/UA experiments, visualize (for better communication of) the results, and reach to conclusions. However, I have some major concerns about some of the methods and tools used in this study that I explain below:

a) Discrete PAWN SA: My most fundamental concern is related the way the main SA with PAWN is performed in this work, which also led to main conclusions in the paper. I strongly feel that the PAWN SA results (Figure 2) is largely impacted by the NUMBER of discrete realizations in each category (Table 4) and not by their CONTENT. In other words, it is intuitive that in this design of SA experiments, by default, the category with a higher number of members will always show a higher influence, because parameters sampled here will naturally have a much higher variability with respect to those categories. And this is exactly what we see in SA results and why some results are rather counter intuitive (e.g. negligible or small influence from land use changes or model setup, and very large influence from Climate and parameters). This is a fundamental issue that needs to be addressed by authors as it is the foundation for all conclusions.

b) Design of Experiments: Authors do a great job particularly in explaining a rather careful and detailed procedure to setup the model, process the required data, define HRUs, and layout future land, pollution, and climate scenarios. This is extensive amount of work. However, I feel that this breadth has caused insufficient scientific depth in places in the manuscript. For example, it is unclear to me why certain various metrics are chosen in the SA analysis with VARS? How are these metrics really differ-

ent from each other from an SA perspective (in particular, NSE and RSR are directly related, so why both are used?), Why this choice is not consistent with the metrics used in the next steps (e.g. what happened to KGE or RSR)? Perhaps strategically reducing some of the metrics can help in a more efficient way of conducting SA and presenting its results (e.g. some of the quantile classes presented in Figure 2 in each signature measure can be removed).

Or for example, what is the scientific reference or justification for the way UA is conducted here at the end using 7K simulations out of all possible combinations? Wouldn't a Latin Hypercube Sampling be a more effective choice than random sampling? These methods and choices (and other similar ones) must to be clearly justified in the manuscript.

5. Are the results sufficient to support the interpretations and conclusions? Please see my comments above in NO.4.

6. Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)? No. Details of SA/UA experiments are missing. In particular, I found description of the VARS method somewhat short and there are important details that are missing (a more careful description from the original papers or some of newer applications is recommended). Another very important information that is missing is the ranges used for parameters, and an explanation of how these ranges are determined. These ranges can impact all the SA/UA results. Or it is unclear how parameters are tied to HRUs, and how all different setups, with different NO. of HRUs, in different basins have the same number of parameters (42) when doing SA with VARS?

7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution? Yes.

8. Does the title clearly reflect the contents of the paper? Yes for the most part.

9. Does the abstract provide a concise and complete summary? Yes.

10. Is the overall presentation well-structured and clear? Yes for the most part.

11. Is the language fluent and precise? I feel the language needs to be modified a bit. Both in terms of English grammar (double check usage of "the" and "comma"), and in terms of being scientifically more precise (e.g. using "pollution" instead of "emission"; or using "most influential input" instead of "most relevant"; or page 3 line 4; or page 3 line 26). I recommend a more careful re-view of the manuscript in this regard.

12. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used? Yes.

13. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated? Some of the quantile classes presented in Figure 2 in each signature measure can be removed.

14. Are the number and quality of references appropriate? Yes.

15. Is the amount and quality of supplementary material appropriate? Yes.